Object colours, material properties and animal signals

Article (Published Version)

© 2019. Published by The Company of Biologists Ltd | Journal of Experimental Biology (2019) 222, jeb204487. doi:10.1242/jeb.204487

RESEARCH ARTICLE

Object colours, material properties and animal signals

Lucas Wilkins[1,2,*] and Daniel Osorio[2]

ABSTRACT

Humans and other animals often use colour to recognise objects regardless of their context – as a measure of material properties rather than of their contrast with a background. Most work on visual communication signals is, however, concerned with colour differences, typically scaled by just noticeable differences (JNDs). Here, we move from the prevailing physiological framework to understand what a given colour or type of colour might tell an animal about an object. To this end, we consider the properties of object colour solids, which represent the colour gamut of reflective materials for a given type of animal eye. The geometry of colour solids reveals general relationships between colours and object properties, which can explain why certain colours are significant to animals, and hence evolve as signals. We define a measure of colour vividness, such that points on the surface are maximally vivid and the 'grey' centre is minimally vivid. We show that a vivid colour for one animal is likely to be vivid for others, and highly vivid colours are less easily mimicked than less vivid colours. Furthermore, vivid colours such as black, white, red and blue, as well as pale colours and certain unsaturated shades, are produced by pure or orderly materials. Such materials are created and maintained against entropic processes. Vivid colours are therefore indicative of ecological affordance or biological function, so it is valuable to have low-level psychological biases towards these colours regardless of any specific significance they might have to the receiver.

KEY WORDS: Colour theory, Colour, Colour solid, Receiver psychology

INTRODUCTION

Animal and plant communication signals carry varied messages: some are attractive while others are defensive; courtship displays appeal to naïve viewers, whereas flowers and aposematic signals need to be memorable. Despite their various functions, and the diversity of colour vision in their natural receivers, signals generally include a limited palette of colours including, black, white, saturated hues, and pale but unsaturated colours such as pink, whereas greys and light browns are infrequent.

Any signal must attract attention and engender a response – it will be ineffective if it is overlooked or ignored. Why then should certain colours attract the attention of an eye with any given set of spectral photoreceptors, and why are the same types of spectra appropriate for animals with different types of colour vison? We show here how answers to these questions might be found if a colour is regarded not simply as a measure of the spectral composition of a light, but rather as a property of physical objects. This entails quantifying colour in new ways and, here, the dropping of two common practices: splitting of colour into independent chromatic and achromatic components and, the scaling of colour distances using discriminability.

Most work on biological signalling treats animal colour vision as a means to discriminate between spectra, and is primarily concerned with the magnitudes of colour differences as measured by just noticeable differences (JNDs) (Kelber et al., 2003; Kemp et al., 2015; Olsson et al., 2017). Accordingly, one might predict that a colour signal will attract a receiver's attention when it differs strongly from the background (Gittleman and Harvey, 1980) or when the pattern itself has a high contrast (Rowe and Guilford, 1996; Aronsson and Gamberale-Stille, 2008).

In its everyday use, 'colour' is not a relative or relational term like 'contrast', but is absolute, and part and parcel of object recognition. We do not say 'the tennis ball is more yellow than the court', but 'the tennis ball is yellow'. The fraction of light reflected from a surface depends upon its chemical composition and (nanoscale) physical structure. Colour vision yields information about these properties. We speak of red faces, blue tea-mugs and so forth, and other animals may be similar. It follows that animals could find certain colours significant because they are characteristic of particular kinds of object.

The significance of a colour might be related to the specific colouration mechanism – for example, if a pigment is costly to produce (Olson and Owens, 1998) – or to associations with particular beneficial or harmful objects (Endler and McLellan, 1988; Endler and Basolo, 1998; Palmer and Schloss, 2010). In contrast and complementary to these approaches, we look here more broadly at the relationship between the composition of materials, their reflectances and their colours. Perhaps surprisingly, we find that there are aspects of object colour that are consistent between different observers and which can be linked to underlying physical properties that are relevant to the psychology of an organism. In particular, we analyse a property of colour we term 'vividness'. Highly vivid colours to the eye of one organism are highly vivid for the eye of any other with the same or a greater number of photoreceptor types (rules 1 and 2 below). We then show that highly vivid colours are especially informative, in that they correspond to fewer materials, and that those materials will be purer and more ordered than less vivid materials. From this, we argue that psychological salience of these colours should be evolutionarily adaptive.

MATERIALS AND METHODS

Numerical solutions

Calculating vividness using linear programming

Finding the spectrum that lies on the boundary in a given direction from the centre can be understood as a linear programming problem. Linear programming is the name given to the optimisation of problems that have a linear objective function and linear constraints.

[1]Department of Zoology, University of Oxford, Oxford OX1 3SZ, UK. [2]School of Life Sciences, University of Sussex, Brighton BN1 9QG, UK.

*Author for correspondence (lucas.wilkins@zoo.ox.ac.uk)

L.W., 0000-0002-0149-3287

Received 6 April 2019; Accepted 26 September 2019

Journal of Experimental Biology (2019) 222, jeb204487. doi:10.1242/jeb.204487

The canonical form used for finding numerical solutions of such problems is:

$$\begin{array}{ll} \text{Minimise} & \mathbf{c}\cdot\mathbf{x} \\ \text{such that} & \mathbf{A}\mathbf{x}\leq\mathbf{b} \\ & \mathbf{D}\mathbf{x}=\mathbf{e}, \end{array} \tag{1}$$

where $\mathbf{x}$ is the vector to be determined, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{e}$ are vector and $\mathbf{A}$ and $\mathbf{D}$ are matrix parameters. The equality constraints ($\mathbf{Dx}=\mathbf{e}$) are usually explicit in numerical procedures, but in mathematical treatments they are usually understood as a pair of inequality constraints: $\mathbf{Dx}\leq\mathbf{e}$ and $-\mathbf{Dx}\leq-\mathbf{e}$. Our task is to put the problem of finding the distance to the boundary into a form that can be understood by standard linear programming routines, such as the simplex method. In this problem, we find the boundary colour by first finding the spectrum which lies on the boundary of solid. We must work with spectra, not colours, as the colour solid is determined by a constraint on the form of the spectrum.

Inequality constraints (form of Ax≤b)

If $\mathbf{x}$ is a vector representing a reflection spectrum, then we have constraints of the form $0\leq\mathbf{x}_i\leq1$, which we can write in block matrix form as:

$$\begin{pmatrix}\mathbf{I}\\-\mathbf{I}\end{pmatrix}\mathbf{x}\leq\begin{pmatrix}\mathbf{1}\\\mathbf{0}\end{pmatrix}. \tag{2}$$

Equality constraints (form of Dx=e)

There are also equality constraints. The colour of the boundary spectrum must lie on a line passing through the reference colour and the centre. The colours ($\mathbf{s}$) lying on this line are given by:

$$\mathbf{s}=k(\mathbf{s}_{\text{ref}}-½)+½, \tag{3}$$

where $\mathbf{s}_{\text{ref}}$ is the reference colour and k is a variable giving the position along the line.

This has the form $\mathbf{s}=k\mathbf{p}+\mathbf{q}$, where $\mathbf{p}=\mathbf{s}_{\text{ref}}-½$ and $\mathbf{q}=(½,½\ldots)$ are vectors. To get this in the form of $\mathbf{Dx}=\mathbf{e}$ that is needed to specify the linear programming problem, we need to do a little algebraic manipulation. We can write the constraint of the colour to a line in the standard form: $\mathbf{Ms}=\mathbf{v}$, where $\mathbf{M}$ is an n-by-$(n-1)$ matrix and $\mathbf{v}$ is an $n-1$ vector. When $\mathbf{F}$ is a matrix containing the discretised relative quantum yield functions (see Eqn 10), the colour $\mathbf{s}$ is given by $\mathbf{Fx}$ and we have $\mathbf{MFx}=\mathbf{v}$, meaning that $\mathbf{D}=\mathbf{MF}$ and $\mathbf{e}=\mathbf{v}$.

There are multiple choices of $\mathbf{M}$ and $\mathbf{v}$ which constrain the colour in exactly the same way. One way of obtaining $\mathbf{M}$ and $\mathbf{v}$ is by picking one dimension for which $\mathbf{p}$ is non-zero (the 'pivot'), using this to solve for k and substituting the solution back. The result of this procedure is an n-by-$(n-1)$ matrix and an $n-1$ vector. The matrix can be understood as having ones on the leading diagonal and $-\dfrac{\mathbf{p}_i}{\mathbf{p}_{\text{pivot}}}$ in the pivot column, with the pivot row removed. For example, for a tetrachromat pivoting with $i=2$, we have, the 4-by-3 matrix:

$$\mathbf{M}=\begin{pmatrix}1 & 0 & -\dfrac{\mathbf{p}_0}{\mathbf{p}_2} & 0\\ 0 & 1 & -\dfrac{\mathbf{p}_1}{\mathbf{p}_2} & 0\\ 0 & 0 & -\dfrac{\mathbf{p}_3}{\mathbf{p}_2} & 1\end{pmatrix}\leftarrow\text{row removed}. \tag{4}$$

The corresponding vector has values of the form $\mathbf{q}_i-\left(\dfrac{\mathbf{q}_2}{\mathbf{p}_2}\right)\mathbf{p}_i$ and, like $\mathbf{M}$, has the pivot row removed:

$$\mathbf{v}=\begin{pmatrix}\mathbf{q}_0-\left(\dfrac{\mathbf{q}_2}{\mathbf{p}_2}\right)\mathbf{p}_0\\ \mathbf{q}_1-\left(\dfrac{\mathbf{q}_2}{\mathbf{p}_2}\right)\mathbf{p}_1\\ \mathbf{q}_3-\left(\dfrac{\mathbf{q}_2}{\mathbf{p}_2}\right)\mathbf{p}_3\end{pmatrix}\leftarrow\text{row removed}. \tag{5}$$

Objective function (form of c·x)

Lastly, we need the objective function which is to be maximised. This measures how far we are from the centre of the solid. A dot product using colour is the choice because it is signed (unlike the Euclidean distance), and is linear. The function

$$(\mathbf{s}-½)\cdot(\mathbf{s}_{\text{ref}}-½) \tag{6}$$

is the distance from the centre, projected in the direction of $\mathbf{s}_{\text{ref}}$. To get this in the form $\mathbf{c}\cdot\mathbf{x}$ we rewrite it as:

$$\mathbf{s}\cdot(\mathbf{s}_{\text{ref}}-½)+½\sum_i(\mathbf{s}_{\text{ref},i}-½). \tag{7}$$

The second term is a constant. This means that its value does not effect the minimisation and we can ignore it, giving:

$$\mathbf{s}\cdot(\mathbf{s}_{\text{ref}}-½). \tag{8}$$

After substituting $\mathbf{Fx}$ for $\mathbf{s}$, and rearranging using the commutativity of the dot product we find that

$$-\mathbf{c}=(\mathbf{s}_{\text{ref}}-½)\mathbf{F}. \tag{9}$$

The minus sign is present because the problem is specified as a minimisation, not a maximisation.

Colour solid geometry

The geometry of the colour solids have been calculated here using a convex pruning method. Through convexity it is possible to demonstrate that the boundary of the colour solid is formed by reflectances that contain only zeros and ones. Every spectrum can be formed though a convex combination of zero-one spectra, thus, any colour must be a convex combination of the colours formed by zero-one spectra. To calculate the solid numerically, we use discretised relative quantum yield functions, $g_i(j)$, for a set of wavelengths λ_0, λ_1, …λ_j…λ_n:

$$g_i(j)=\int_{\lambda_j}^{\lambda_{j+1}}f_i(\lambda)d\lambda, \tag{10}$$

where i indexes the photoreceptor class, and j indexes the wavelength band. The algorithm works by incrementally enumerating all the reflectance spectra that may lie on the surface of the solid, ignoring all those for which this is clearly impossible.

Denote the set of colours defined by non-zero discretised reflectance only in the region $[\lambda_0,\ \lambda_k]$ by $\mathcal{S}_k$. This will be uniquely defined by a finite set of points lying on the convex hull of that set, $\mathcal{H}_k$. The sets $\mathcal{S}_k$ and $\mathcal{S}_{k-1}$ are related to each other by:

$$\mathcal{S}_k=\mathcal{S}_k\cup\{\mathbf{s}+(g_1(k),g_2(k)\ldots)|\mathbf{s}\in\mathcal{S}_{k-1}\}. \tag{11}$$

Journal of Experimental Biology (2019) 222, jeb204487. doi:10.1242/jeb.204487

The first and second terms of the right-hand sum correspond to the addition of reflectances in region $[\lambda_{k-1}, \lambda_k]$ with values 0 and 1, respectively. Thus $\mathcal{S}_n$ contains all colours corresponding to discretised reflectances with zero or one everywhere, and has 2^n elements. For resolutions as low as n=10 the full enumeration of this can be problematic. The effect this large set has will only be made worse by the computational complexity convex hull algorithms, which is super-linear in both the number of points and the dimension. To avoid having such large sets we prune away fruitless candidates as we go along.

For calculating the bounding points of the $\mathcal{S}_k$, only the values $\mathcal{H}_{k-1}$ matter. This is because the addition of a vector, $\mathbf{v}$, to any point in $\mathcal{S}_k$ will result in a point that is a convex combination of points in $\{\mathbf{s}+\mathbf{v}|\mathbf{s}\in\mathcal{H}_k\}$. If we let $\mathcal{C}$ denote the operation which obtains the points defining the convex hull, then it follows that:

$$\mathcal{H}_k = \mathcal{C}(\mathcal{H}_k \cup \{\mathbf{s} + (g_1(k), g_2(k)\ldots)|\mathbf{s}\in\mathcal{H}_{k-1}\}), \quad (12)$$

which, when applied iteratively, results in a computationally feasible number of points. The speed of this procedure may be increased by including the results of a low-resolution version of the calculation (even just the white point). Call these known points $\mathcal{K}$, and then instead iterate using:

$$\mathcal{H}'_k = \mathcal{C}(\mathcal{K} \cup \mathcal{H}'_k \cup \{\mathbf{s} + (g_1(k), g_2(k)\ldots)|\mathbf{s}\in\mathcal{H}'_{k-1}\})\backslash\mathcal{K}, \quad (13)$$

with:

$$\mathcal{H}_n = \mathcal{C}(\mathcal{H}'_n \cup \mathcal{K}). \quad (14)$$

The last application of $\mathcal{C}$ may be omitted if we know that $\mathcal{K}\subseteq\mathcal{H}_n$ – a condition which is desirable for the efficiency of the pruning. The computational complexity of this algorithm depends on the number of photoreceptor classes and becomes difficult for observers with more than 5 photoreceptors.

Using linear programming

A linear programming method like the one described in the vividness calculation can be used to obtain a collection of points describing the geometry of the colour solid. To do this, one samples directions around the centre, and calculates the corresponding boundary spectra using the linear programming approach described above. This method has a disadvantage in comparison with the method below in that it can only resolve fine detail such as points or sharp edges by using a very high density of points.

Metameric colours

Either of these algorithms can be used to calculate the colours for one observer that correspond to the colours for another. Call the discretised relative quantum yield functions for the two observers $a_i(j)$ and $b_i(j)$, respectively, and let $c_i(j) = (a_0 \ldots a_{m_a}, b_0 \ldots b_{m_b})$, where m_a and m_b are the number of photoreceptor classes for observers a and b. An m_a+m_b dimensional colour solid can then be constructed for c_i. The set of observer b colours corresponding to an observer a colour can then be calculated as the m_b dimensional cross-section of the m_a+m_b dimensional solid for which the observer a coordinates are constant. As this involves calculating a m_a+m_b dimensional colour solid, it is only sufficiently fast on current machines for m_a+m_b around 5 and below. For other cases, a linear programming approach is prudent.

RESULTS

Object colour solids and vividness

Objects are seen by reflected light, and the gamut of all possible reflection spectra can be represented by their locations within a Cartesian space known as an object colour solid whose axes correspond to photoreceptor excitations relative to the illumination spectrum (Figs 1 and 2; Vorobyev, 2003, Koenderink, 2010). Colour solids include the three main aspects of colour, namely hue, saturation and brightness, whereas the more familiar chromaticity diagrams, such as Maxwell's triangle (Maxwell, 1860) discount brightness. An additional difference lies in the nature of the gamut boundaries. In a chromaticity diagram, the boundary is defined by monochromatic spectra and (for trichromats) the purple line. As a monochromatic reflection contains negligible light, highly saturated (or pure) colours are dark. By comparison, object colour solid boundaries include both black (zero reflection) at the origin and white (maximal reflection at all visible wavelengths), with the intermediate boundary surface being well approximated by spectral step functions (Fig. 2). Such spectra can be bright and are more representative of natural spectra than monochromatic spectra.

Following the logic that colour refers primarily to the physical properties of reflective materials (or objects), we propose a measure of colour within the object colour solid, which we call 'vividness'. Vividness resembles colorimetric parameters such as purity or saturation, but achromatic colours – black and white – and pale, unsaturated colours can be highly vivid. We show mathematically and empirically that the vividness of reflectance spectra is well correlated between different types of colour vision. Consideration of the relationship between vividness and the physical processes that generate object colour demonstrates that orderly nanostructures or pure pigments are typically more vivid than their less orderly or pure materials. As order emerges against entropic tendencies, vivid or 'bright' (Hamilton and Zuk, 1982) colouration is indicative of a functional role, and therefore more 'meaningful', and hence such colours worthy of greater attention. We start with an account of why the achromatic colours black and white are maximally vivid, leading to a general model which includes chromatic colours.

Black and white

Models of colour vision and colour appearance usually treat chromaticity, which combines hue and saturation, as qualitatively distinct from the achromatic component of colour. This distinction is grounded in physiology and psychophysics (Livingstone and Hubel, 1988; Osorio and Vorobyev, 2005), yet black and white are 'colours' in ordinary English usage, and they are common in biological signals. In physiological terms, an animal sees black when its spectral photoreceptors all have a low excitation, and white when they all have a high excitation. Hence, spectra that look black have low intensity at all visible wavelengths and spectra that look white have high intensity. As there is only one perfectly black spectrum, and a perfectly white surface must reflect maximally at all wavelengths it follows that spectra responsible for black and white are the same for all observers that share the same range of visible wavelengths, with minimal or maximal excitation across that range. These spectra can be designated as 'extreme', because they are at the limits of receptor excitation achievable by a surface.

Unlike black and white, intermediate reflectance spectra can produce different excitations according to the particular set of photoreceptors in a given eye. For the simplest case of two eyes each with a single type of photoreceptor, but tuned to different wavelengths, a stimulus of intermediate intensity can give different receptor responses in the two monochromatic observers.

Journal of Experimental Biology (2019) 222, jeb204487. doi:10.1242/jeb.204487

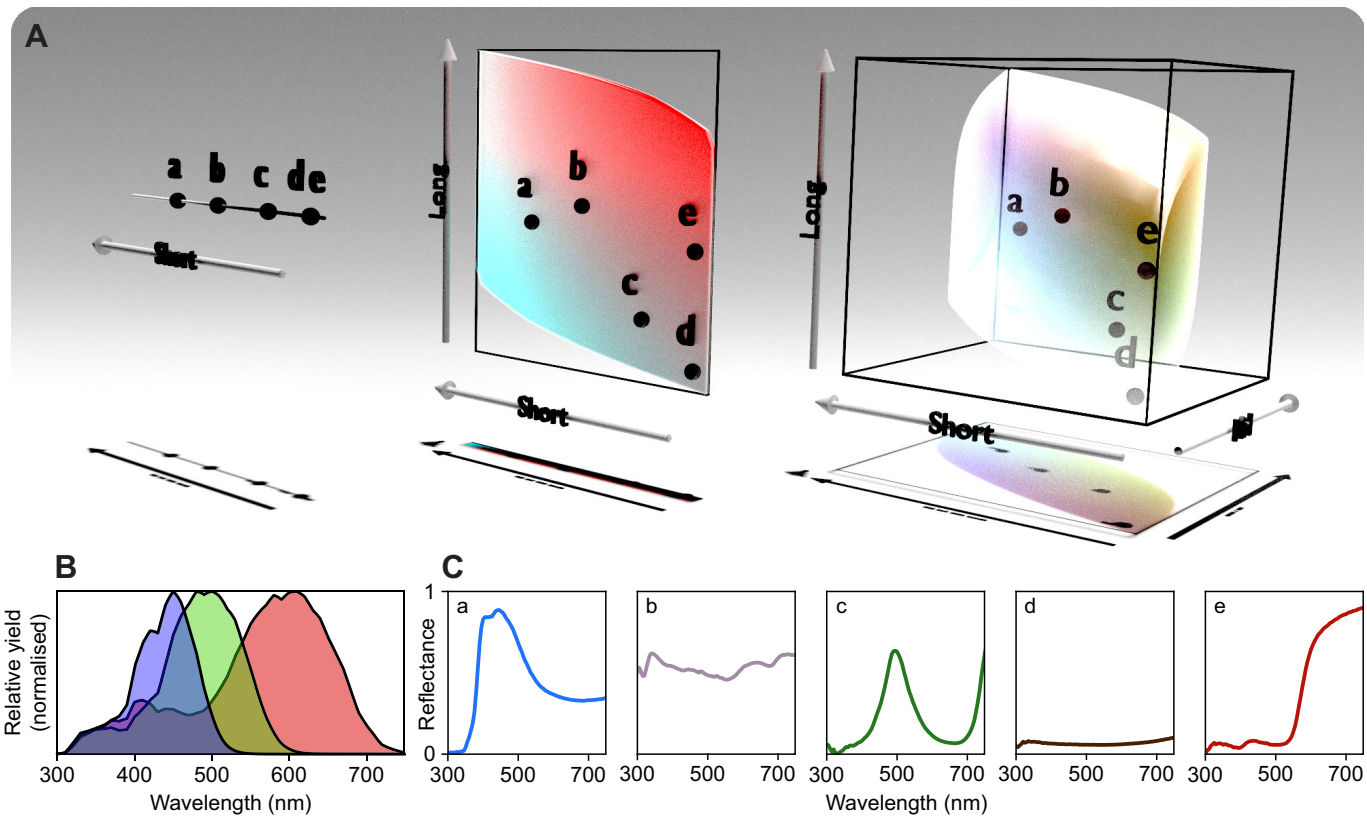

Fig. 1. The object colour solid. (A) Colour solids for a monochromat (left), a dichromat (middle) and a trichromat (right). Note how the positions of colours (a–e) in the pre-existing dimensions do not change as more photoreceptor classes are added. (B) Relative quantum yield curves (spectral sensitivity multiplied by the illuminant and normalised). The monochromat has just the short wavelength sensitive photoreceptor class, the dichromat contains the short- and long-wavelength-sensitive photoreceptor classes, the trichromat contains all of them. Each 'expands' the solid outwards into the new dimension. (C) The natural reflectance spectra responsible for colours (a–e).

Consequently the appearance of greys for different types of monochromat is less predictable than are black and white. Similarly, for any given monochromatic eye only one spectrum can give black or white, but many different spectra can produce indistinguishable intermediate responses, a phenomenon which is known as metamerism (Logvinenko, 2009). Black and white, produced by extreme spectra, are maximally vivid colours, while greys have lower vividness.

Chromatic components of colour

The best known geometric representations of colour are chromaticity diagrams in which desaturated colours lie near the centre and the most saturated colours at the extremities, with the angle around the centre specifying hue. One such diagram is Maxwell's triangle. Using a method based on colour mixing, Maxwell specified the colours of monochromatic lights as mixtures of red, green and blue primaries. When the brightness is ignored these results can be drawn in a triangle, with pure primaries at each corner, and mixtures within. The line of all the monochromatic lights, the 'monochromatic locus' or 'spectral line' forms a rounded Λ shape along two of the edges, with the ends joined by a 'purple line'. The fractional distance from the centre to the edge is then a measure of spectral purity, which is related to the colour's perceived saturation (Wyszecki and Stiles, 2000).

For colour vision of non-human species, similar diagrams are typically based on photoreceptor spectral sensitivities (Renoult et al., 2017). For an eye with n spectral types of photoreceptor (contributing independently to colour vision) the chromaticity diagram is $n-1$ dimensional. This kind of chromatic space is good for describing the colours of lights, but less suitable for reflectance spectra, firstly because there are very few (if any) narrow, monochromatic reflectances and secondly, because a reflectance can only approach the boundary by reducing the amount of reflected light, making the colour darker. Maximal spectral purity is black and hence chromaticity diagrams typically exaggerate differences between dark colours.

Object colour solids are so-named because they appropriate for representing 'object' or reflectance spectra. They are useful in colour reproduction and the formulation of dyes and pigments, because the available gamut can be compared with the colour range visible to the human eye. The axes of colour solids (Wyszecki and Stiles, 2000; Schrödinger, 1920; Vorobyev, 2003; Koenderink, 2010) correspond to photoreceptor excitations (or a similar set of primaries) normalised to the illumination intensity. As photoreceptor spectral sensitivities overlap, colour solids do not fill the space defined by the axes, but are roughly ellipsoidal with two pointed corners (Fig. 1). Monochromatic spectra lie an infinitesimal distance from the origin which is black, and maximal reflectance (white) is at the opposite vertex. Humans have three types of cone photoreceptor, and hence a 3D object colour space, but the same geometrical principles apply to any type of colour vision: for most mammals, which are dichromats, the space is 2D, whereas the spaces of birds are probably 4- or 5-dimensional

Journal of Experimental Biology (2019) 222, jeb204487. doi:10.1242/jeb.204487

(Kelber et al., 2003; Vorobyev, 2003). As there are both empirical and mathematical relationships between these spaces, we refer to all of them as colour solids.

Object colour solids

We now more formally define the colour solid and vividness. For a given set of photoreceptor spectral sensitivities [$s_i(\lambda)$, $i \in 1 \ldots n$] and an illuminant [$l(\lambda)$], reflectance spectra [$r(\lambda)$] can be organised into a geometric object known as the object colour solid (Schrödinger, 1920; Wyszecki and Stiles, 2000), shown in Fig. 1. This object is formed from the colours [a vector of photoreceptor quantum yields, $(q_1 \ldots q_n)$] associated with all theoretically possible reflectance spectra: that is all distributions of reflectance values between zero and one, taken over visible wavelengths (Λ):

$$q_i = \int_{\Lambda} s_i(\lambda) l(\lambda) r(\lambda) d\lambda. \quad (15)$$

To obtain the colour object solid, photoreceptor responses (q_i) are normalised to the quantum yield of a perfectly reflecting surface:

$$x_i = q_i / q_i^{\max} \quad \text{where} \quad q_i^{\max} = \int_{\Lambda} s_i(\lambda) l(\lambda) d\lambda. \quad (16)$$

This means that black has x coordinates of (0, 0, …) and white has coordinates (1, 1, …). The maximal reflectance could instead be set according to a standard such as $BaSO_4$ pellet – this does not affect the arguments we make here and should in fact result in a more practical measure of vividness. Throughout, we express the x coordinates in terms of a relative quantum yield function:

$$x_i = \int_{\Lambda} f_i(\lambda) r(\lambda) d\lambda. \quad (17)$$

The object colour solid is convex, centrally symmetric, and lies within a unit n-cube (Fig. 1). It is pointed at the diagonally opposite black and white corners, but the other corners and edges are smoothed due to the spectral overlap of the photoreceptors. Except for the boundaries, every point in the solid maps to more than one reflectance spectrum, corresponding to the phenomenon of colour metamerism (Logvinenko, 2009). The boundaries represent reflectances with the highest spectral purity for a given luminance. Unlike chromaticity diagrams, the colour solid therefore accounts for the trade-off of saturation against luminance, so light colours can lie on the boundary. This accords with the intuition that we do not see light colours as necessarily less pure than dark.

The exact calculation of the colour solid is rather involved. Numerical solutions can be obtained by dynamic programming (Wyszecki and Stiles, 2000) or our own method (see Materials and Methods for details of both). For the purpose of our argument, especially rule 1 below, it is better to start with Schrödinger's approximation (Schrödinger, 1920; Vorobyev, 2003), which uses spectra formed by step changes in intensity between zero and one. Fig. 2 illustrates the Schrödinger's spectra for dichromats. The number of step changes depends on the number of spectral classes of photoreceptor: for an eye with n spectral classes the maximum number of step changes needed to approximate the colour solid boundary is $n-1$. For dichromats, they are single steps, the two series (step up and step down) ranging from black through red, orange, then yellow to white, and from black through blue then cyan to white. For trichromats there are two steps in the visual range, while for a tetrachromat the Schrödinger spectra include those that have a reflectance of 0 up to a wavelength λ_1 then 1 until another, λ_2, then 0 until λ_3 then 1, and in addition their inversions, i.e. those that have values of 1 then 0 then 1 then 0. Whilst the Schrödinger spectra are only approximations to the extreme spectra that lie on the boundary of the solid, the extreme spectra, like Schrödinger spectra, will only ever have either maximal (1) and minimal (0) intensity at every wavelength. When discretised approximations like those in the Materials and Methods are used to calculate boundary spectra, a step transition between two discretisation points will appear as an intermediate intensity for the corresponding section.

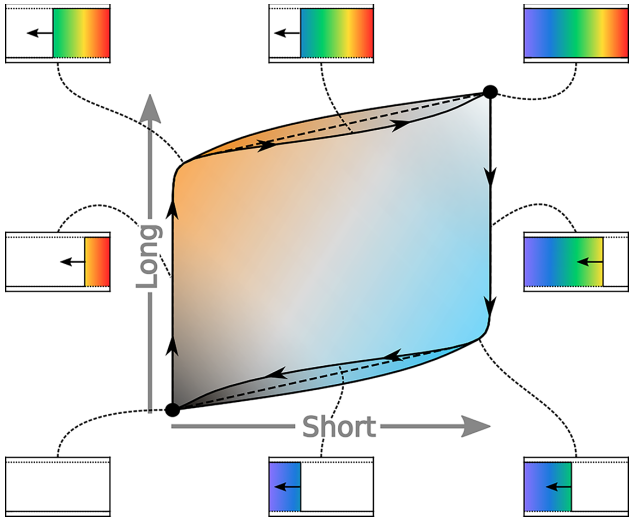

Fig. 2. Calculation of the Schrödinger approximation for the dichromat observer in Fig. 1. With dichromats, the Schrödinger approximation can be thought of as tracing out a closed curve made of spectra with a single step change. On the left and upper portions of the solid, the spectra have a value of 1 at long wavelengths and 0 at short wavelengths. As transition point moves from long to short, the colour moves from black to white. On the right and lower portions of the solid, the spectra have a value of 0 at long wavelengths and 1 at short wavelengths. As the transition point moves from long to short, the colour moves from white to black. The Schrödinger approximation does not fully cover the whole colour solid. In this case, the Schrödinger approximation gives a concave which does not perfectly describe the colour solid, and instead lies slightly within it. The convex hull of the Schrödinger approximation (dotted lines) is closer, but still does not quite match. For this observer, the dichromatic Schrödinger spectra have a mean vividness of ≈0.943 (all in the range 0.843 to 1). These failures of the approximations are most unlikely to have any consequence for the consideration of the vividness of natural colours. The colours used to render this solid are only indicative of their true appearance.

Schrödinger's approximation of boundaries of the object colour solid holds best when photoreceptor spectral sensitivities have a single peak. Some receptor spectral sensitivities are bimodal (e.g. Fig. 4), in which case, Schrödinger's spectra typically lie slightly inside the boundaries of the corresponding colour solids. Such deviations are in practice small, because photoreceptor spectral sensitivities are typically smooth with a dominant peak. To the extent that this approximation holds a conclusion analogous to that for black and white applies; namely that for eyes with the same number of photoreceptor classes, the Schrödinger spectra corresponding to boundary colours are the same. Differences between receptor sensitivities mean that the shape of the object colour solid varies between species, and the locations of individual spectra within the boundary are not directly comparable (Fig. 4), but they will be on the boundary nonetheless. It follows that Schrödinger spectra for one eye are Schrödinger spectra for an eye with a larger number of photoreceptor classes: the set of

Journal of Experimental Biology (2019) 222, jeb204487. doi:10.1242/jeb.204487

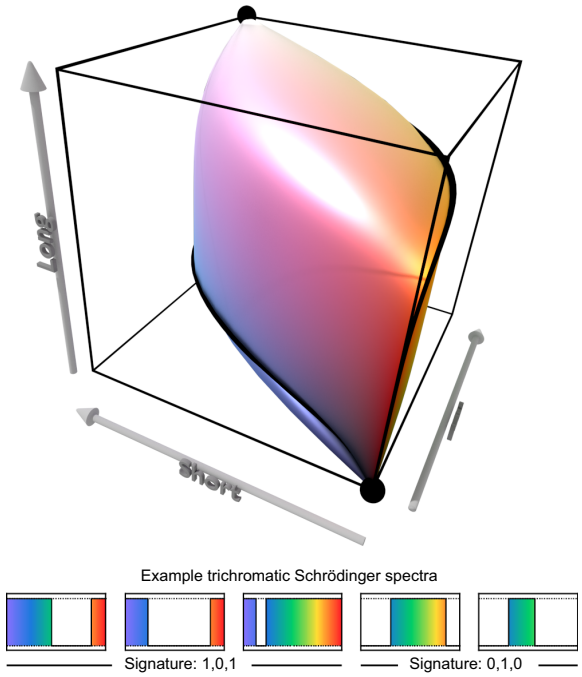

Fig. 3. The Schrödinger approximation for the trichromat observer in Fig. 1. The coloured surface produced by the trichromatic Schrödinger spectra are characterised by having two step changes. The black ribbon running around the edge of the solid shows the colours corresponding to the dichromat Schrödinger spectra used in Fig. 2. Note that they do not lie in a plane. The spheres at the tips correspond to the monochromat Schrödinger spectra (i.e. black and white). Just as the dichromat Schrödinger approximation is formed by two curves lying between black and white (one with reflectance of 0 at long wavelengths which transitions to 1 and the other with 1 at long wavelengths transitioning to 0), the trichromat Schrödinger approximation is formed from two surfaces, one with spectra with a 'signature' of 0,1,0 and the other with 1,0,1. The two surfaces meet at the curves determined by the dichromat Schrödinger spectra, which themselves meet at the points determined monochromat Schrödinger spectra. This is representative of a general pattern that applies to all observers. As in Fig. 2, the surface formed by Schrödinger spectra is concave in some places. This concavity is most easily observed in the purple coloured region on the bottom left as viewed. The colours used to render this solid are only indicative of their true appearance.

Schrödinger spectra for an n-chromat contains the Schrödinger spectra for an m-chromat, when $m \leq n$.

The tendency for the same spectra to lie at the boundary of the colour solid extends to colours not on the boundary, although, as was the case for black, white and grey, the variability in the position is greater for more central colours (Fig. 4). Whilst the theoretical limit of this variability, illustrated in Fig. 4, is relatively high, the spectra that achieve this limit (the boundary of the metamer set) are difficult to realise practically. This is evident in Fig. 5.

In practice, spectra of natural objects are a small subset of physically possible spectra (Maloney, 1986; Osorio and Bossomaier, 1992; Vorobyev et al., 1997), and natural spectra tend to be smoother than Schrödinger spectra (Maloney, 1986) (compare Figs 1 and 2 with Fig. 3). Also, spectra with multiple transitions within the visible range are unusual, so that higher order Schrödinger spectra are seldom approached in nature; exceptions include some structural colours, such as that found on the nape of the feral pigeon, and which roughly approximates a tetrachromat Schrödinger spectrum. Whilst this can appear greyish to trichromats, it is likely to be vivid for birds (Osorio and Ham, 2002).

Definition of vividness

We define vividness as a number ranging from zero at the centre of the colour solid to one at the boundary, which is linear with respect to the colour solid coordinates (x_i). For an observer/illumination combination that is described by n relative quantum yield functions ($f_1 \ldots f_n$) the vividness of a reflectance r is:

$$V(r;f_1 \ldots f_n) = \frac{\|\mathbf{x} - ½\|}{\|\mathbf{b}(\mathbf{x}) - ½\|,} \tag{18}$$

where $\mathbf{b}(\mathbf{x})$ is the position of the boundary in the direction of the vector $\mathbf{x} - ½$ from the centre $(½, ½ \ldots)$. $\| \cdot \|$denotes a vector's length.

The numerator is the Euclidean distance from the centre of the solid. The division by $\|\mathbf{b}(\mathbf{x}) - ½\|$ maintains an invariance between species. It has the effect of adjusting the numerator to reflect more physical, rather than perceptual, properties. The latter being more difficult to assess for non-human species (and even for humans).

Properties of vividness

The properties of colour solids we have discussed so far can be formally expressed with two rules, the first is motivated by our argument above and it is true insofar as it is approximate, and contingent on having 'well behaved' relative quantum yield functions spanning a comparable visual range, the second is a geometric fact:

$$\text{Rule 1 : approximate equality} \quad V(r;f_1 \ldots f_n) \approx V(r;g_1 \ldots g_n). \tag{19}$$

For two observers with equal numbers of spectral receptor classes (n), the vividness of a stimulus is approximately the same for both observers. This rule is motivated by the foregoing discussions, and is corroborated by Figs 4 and 5. For more vivid colours, the range of discrepancy decreases, and the approximation is better – a consequence of there being 'fewer' metamers. This phenomenon can be seen in Fig. 4.

As this rule concerns the relative quantum yield functions f_i and g_i which are obtained from both spectral sensitivities and the illuminant, rule 1 is both a statement about changes in photoreceptor sensitivities and in illumination.

$$\text{Rule 2 : monotonicity with dimensionality} \quad V(r;f_1 \ldots f_n) \leq V(r;f_1 \ldots f_n,f_{n+1}). \tag{20}$$

For a given illumination, an increase in the number of receptor types will result in an increase in vividness, so that colours are more vivid for species that have a large number of photoreceptor classes (e.g. birds) than for those with fewer (e.g. mammals). This effect can be observed in Fig. 5, and a graphical proof is given in Figs S1 and S2. These rules allow one to describe the relationship between any two observers.

As vividness increases the constraints on the variety of spectra that can realise the colour become increasingly restrictive, until, at the boundary of the solid there is a unique spectrum.

This rule can be compared with the metamer mismatching transformation that occurs under variable illumination (Tokunaga and Logvinenko, 2010). Colour purity (Wyszecki and Stiles, 2000) resembles vividness for the chromatic plane, and has fairly similar properties. This is because the chromaticity space is a scaled cross-section of the colour solid at fixed luminance as it goes to zero (i.e. a slice of the solid very near black). Vividness also resembles other measures, such as chroma (Endler, 1990). In addition to the

Journal of Experimental Biology (2019) 222, jeb204487. doi:10.1242/jeb.204487

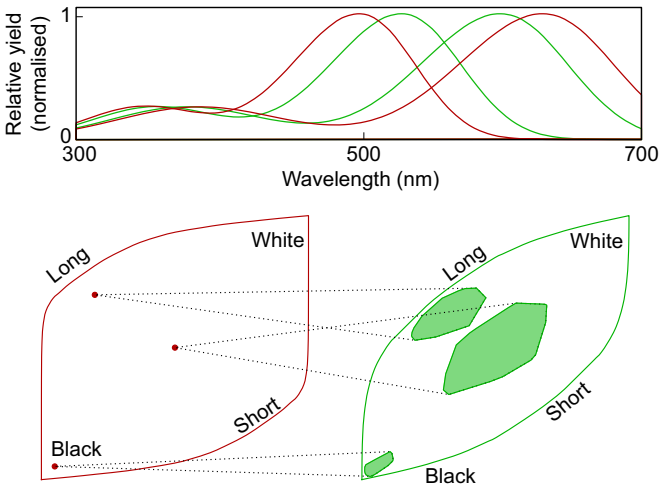

Fig. 4. The relationship between the colours in two dichromat object colour solids. Because of metamerism, points on the left correspond to multiple points (areas) on the right. The exact location depends on the particular underlying spectrum. Similarly (although not shown), points on the right correspond to areas on the left. At the edge of the solid, the correspondence is one to one, and the size of the region in one corresponding to a point in the other increases towards the centre. These mapped areas, like the solids themselves, are based on the assumption that spectra may instantaneously transition between 0 and 1. Such transitions cannot be realized physically, and the spaces and areas are theoretical bounds. The filled areas on the right are bounded by spectra with a greater number of transitions than the solid. These filled areas are conservatively large estimates of the degree to which metamerism is physically realisable. The relative quantum yields were produced from an A1 pigment template which has a pronounced β-peak (Govardovskii et al., 2000), but the correlation in vividness remains. This demonstrates how rule 1 arises without the use of the Schrödinger approximation.

difficulties we have already highlighted with chromaticity, purity is not defined for monochromats, is not mathematically well behaved for dichromats, and can be quite complex for tetrachromats and above.

Mixing

Vividness has a fundamental relationship to physical order, as exemplified by the case of conservative mixing. That is to say, where two colours, $\mathbf{x}_a$ and $\mathbf{x}_b$, are mixed in a ratio of z_a:z_b and the resulting colour is:

$$\mathbf{x}_{\text{mixed}} = \frac{z_a\mathbf{x}_a + z_b\mathbf{x}_b}{z_a + z_b} = k\mathbf{x}_a + (1-k)\mathbf{x}_b \qquad (21)$$

where $k = \dfrac{z_a}{(z_a + z_b)}$ and the various values of k correspond to various positions on the line segment between $\mathbf{x}_a$ and $\mathbf{x}_b$. This holds regardless of the spectra that produce $\mathbf{x}_a$ and $\mathbf{x}_b$. The distance of points on this line segment are closer to (or the same distance from) the centre of the solid than the more distant of $\mathbf{x}_a$ and $\mathbf{x}_b$. If the solid were a perfect sphere (i.e. if the distance to the boundary were fixed) we could conclude directly that vividness of the mixture was smaller than that of the colours being mixed. As the solid is not a sphere, we must make a further observation: where the surface is (strictly) convex, a line segment connecting two points on the surface of the solid passes entirely within its volume (see Figs S1 and S2). The colour solid's boundary colours have the distinguishing feature that they cannot be made by mixing other object colours.

The rule applies to all observers, in spite of the fact that numerical values of V for a particular object colour may well be different.

$$\begin{aligned} &\text{Rule 3 : convexity of mixing} \\ &V(kr_1 + (1-k)r_2; f_1 \ldots f_n) \leq \\ &\quad \max\{V(r_1; f_1 \ldots f_n), V(r_2, f_1 \ldots f_n)\}, \end{aligned} \qquad (22)$$

for all $k \in [0, 1]$. Mixing two colours results in a colour that is less vivid than the most vivid of the two (and perhaps less than both). Rule 3 can be written in a more general form, for a mixture of multiple colours, as:

$$V\left(\sum_i k_i r_i; f_1 \ldots f_n\right) \leq \max\{V(r_i; f_1 \ldots f_n)\}, \qquad (23)$$

with $k_i \in [0, 1]$ and $\sum_i k_i = 1$.

This expression describes a phenomenon familiar to anyone who has mixed paints, once a duller colour is mixed into a more vivid one there is no way to recover the original vividness except by adding an even more vivid paint. It does not, however, correspond to all colour production; an important example being the layering of pigment and nanostructural components used in structural colouration, which is not well modelled by this process.

This is not a property unique to vividness, but we must bear in mind the significance of this rule for our argument: without this rule, we would only have mathematical results about the geometry of colour solids, but with this rule, we can talk about the physical properties of vividly coloured materials. On its own this does not address how the vividness is perceived, but we can tackle this by taking a cue from Gibson's ecological psychology (Gibson, 1979). In the Gibsonian paradigm, the perception of an object derives from its functional properties: if we can talk about physical properties of objects and their relevance to an organism we can speak about (if only indirectly and *ceteris paribus*) the corresponding psychological structure. Although we cannot directly address how an organism perceives vividness, we can ask about the meaning of it in terms of an organism's ecology, and this we would expect to manifest psychologically.

We have found that vividness is comparable for different species and that there are physical processes which act on it in the same way; thus we expect the psychological significance of vividness to be in some way comparable. Although this is not a very detailed conclusion, it is important because it gives us reason to believe that the mathematical results we have presented have a real psychological significance, and an impact on biological processes. It is more speculative to ask about how exactly vividness might manifest psychologically. However, the kinds of physical processes that affect vividness allow us to make some suggestions. We will explore these in closing sections.

DISCUSSION

Object colour solids are useful representations of the colours of reflective surfaces, as opposed to lights. The mathematical properties of these spaces along with empirical evidence (Fig. 5) leads to three main findings relevant to the evolution of colour in communication signals. Firstly, we define a measure of colour, vividness, which corresponds to the distance of a colour from the (grey) centre of the solid, and show mathematically that a spectrum which is highly vivid (i.e. near the boundary of the colour solid) for one type of colour vision will generally be highly vivid for any type

Journal of Experimental Biology (2019) 222, jeb204487. doi:10.1242/jeb.204487

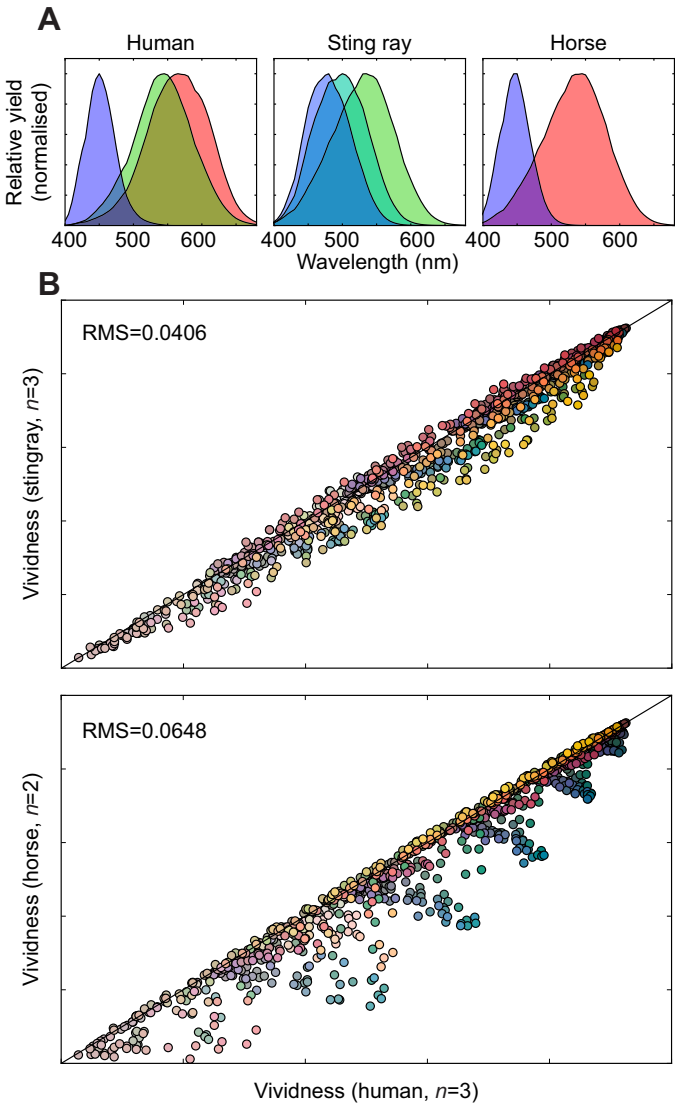

Fig. 5. Correlation of vividness between three disparate species. (A) Example relative quantum yield functions with D65 illumination (Wyszecki and Stiles, 2000) for trichromatic humans (Stockman and Sharpe, 2000) and stingrays (Hart et al., 2004), and dichromatic horses (Carroll et al., 2001). The scatter plots compare the vividness of all 1269 Matt Munsell chips (Munsell et al., 1950; Parkkinen et al., 1989) – a collection of coloured stimuli designed to cover the human colour gamut uniformly – for the three species. (B) Rule 1 (top) shows the approximate equality between vividness for the two species with the same number of cone classes (humans and sting rays). The effects of rule 2 (bottom) indicate that whilst the approximate equality holds for many cases, some chips that we see as blue or pink can be less vivid for a horse. RMS values are the root mean squared errors between the two observers' vividness values over the entire Munsell dataset.

that has the same number or more spectral types of photoreceptor. We find empirically that the vividness of natural reflectance spectra is correlated for eyes with different sets of photoreceptors (Figs 4 and 5). Secondly, colours on the boundary of the solid are attributable to a single reflectance spectrum, and the number of spectra that map to a given point in the colour solid increases towards the centre, so that many spectra can look 'mid-grey'. Highly vivid colours are therefore more likely to be associated with a specific physical cause (i.e. material), and they will be less easily reproduced by alternative means than less vivid colours. Thirdly, mixing colours inevitably reduces the vividness of the more vivid colour, so that pure materials tend to be more vivid than mixtures. This is why vacuum-cleaner dust is greyer than home furnishings. For structural colours increasing regularity of nanostructure increases vividness, and mixing of pigments can only render them less vivid.

Owing to entropy, order does not arise by chance in nature, so vivid colours are indicative of some functional role. This need not be as a signal; the vivid colours of leaves and blood are due to high concentrations of light harvesting and oxygen transport pigments. This is not to say that a dull-coloured tissue cannot have a specific function, the implication only works one way – vividness requires some kind of order, but order does not necessarily result in vivid colour. Nonetheless, if an object has a functional role it is *a priori* worthy of attention, which is a requirement for any signal. Historically, vivid colouration has been associated with the phenonemon of life, and this in turn has been related to thermodynamics. In *Tropical Nature* (Wallace, 1878), Wallace argues that colourfulness is a consequence of 'vital energy', and hence a natural attribute of living organisms. While Schrödinger proposed in his book *What is Life?* (Schrödinger, 1944) that life is characterised by order away from chemical equilibrium – that is, by being non-entropic – here, we have seen why vivid colours are indeed likely to be associated with a system, such as life, that counters the effects of entropy.

Aposematism and mimicry

The strong contrasts and vibrant colours of aposematic displays illustrate some of our findings. Vivid colours are often salient, so they have the potential to promote both innate and learnt responses, and they will be seen consistently by a broad range of viewers, so they need not be predator specific. For instance, in tropical forests, predators of insects include monochromatic strepsirrhine primates, dichromatic mammals and snakes, trichromatic primates and amphibians, and tetrachromatic birds and lizards (Kelber et al., 2003). A consequence of rule 2 is that the achromatic extreme colours – black and white – will be effective for all receivers, those approximating the dichromatic Schrödinger spectra (which we would see as two series black, red, yellow, white and black, blue, cyan, white) would be effective for dichromats and above, the trichromatic series which adds purples and greens for trichromats and above, and so on. Similar principles apply in marine environments where predator colour vision ranges from monochromacy in cephalopods through di-, tri- and tetrachromacy in various fish to the multispectral system of stomatopods (Marshall et al., 2015).

A further benefit of vivid defensive signals arises because as vividness increases, copying a colour entails more-exact matching of the spectrum. A defended aposematic model may in principle outmanoeuvre a Batesian mimic by adopting more-vivid colours outside the mimic's physiological 'gamut' (Franks et al., 2009; Briscoe et al., 2010; Bybee et al., 2011) and conversely there might be a pressure for Müllerian mimics to share less-vivid colours.

Modelling of colour in biological signals

Investigation of colour signalling by animals and plants often starts by modelling photoreceptor responses to reflectance spectra. Receptor responses do not directly specify colour differences or colour appearance, which requires further psychophysical detail (Kelber et al., 2003; Kemp et al., 2015). Models based on chromaticity assume that lightness (or luminance) is discounted, and although some add achromatic contrast as a separate parameter (Siddiqi et al., 2004; Olsson et al., 2017), all such models can lead to difficulties. For example, they predict that dark colours are

unrealistically distinct, and there may be an implicit assumption that the strong achromatic contrasts, which are present in many signals are either irrelevant or have a qualitatively different function from chromatic components.

Object colour solids represent the full gamut of colours visible to an eye, and importantly offer a natural means of representing colour as a property of reflective materials rather than spectral lights. It is straightforward to define the locations of reflectance spectra in the object colour solid, and hence to explore broader questions about the gamuts of colour signals directed at various types of receivers (cf. Osorio and Vorobyev, 2008). Vividness is a simple and well defined measure of colour, within a colour solid which can be related to the physical properties of coloured materials, and gives insight into how and why animals with diverse visual systems might evaluate colour. Of course, it remains an empirical question as to whether vividness is a useful measure of colourfulness. One could test whether vividness should predict attention or salience better than colour saturation or purity, or to a scale based on colour distances measured in terms of JNDs in the colour solid.

Conclusion

In *The Origin of Species*, Darwin writes '…the belief that organic beings have been created beautiful for the delight of man […] has been pronounced as subversive of my whole theory…' (Darwin, 1859), but contemporary literature offers a range of accounts of why certain colours or patterns should be attractive to animals. Some refer to the nature of the sensation, or postulate general aesthetic principles, whereas others highlight the specific value of a stimulus. There are two broad concerns here, one is explaining the appearance and stability of a signal in terms of evolutionary costs and benefits (Fisher, 1930; Zahavi, 1975; Grafen, 1990; Hill, 1991; Johnstone, 1995; Prum, 2012), the other is how signals relate to the behaviour and psychology of an animal – receiver psychology and sensory bias (Guilford and Dawkins, 1991; Endler and Basolo, 1998).

Our discussion of vividness has less to say about costs than about receiver psychology. Whilst there is necessarily some kind of immediate thermodynamic cost to maintaining order, this might be only a tiny contribution to the full evolutionary cost, and could be entirely negated by other benefits. Indeed, colour is often a side effect of having an unrelated functional structure (such as in leaves or blood, as mentioned above).

Sensory bias – the idea that mate preferences are shaped by evolutionary pressures on sensory systems – has a number of proposed mechanisms. One of these is that particular colours occur in courtship displays because they resemble objects of value such as food items (Allen, 1879; Endler and Basolo, 1998); this is very different to our account. In contrast to this kind of sensory exploitation, we claim that sensitivity to vivid colours is a general psychological trait that has evolved in response to the effect of basic physical processes on colour. It addresses the question of why an organism would evolve to be 'stimulated' (Fuller et al., 2005) by colours of a certain kind, and to some clues about how the notion of being stimulated might be refined to have greater psychological detail. This is not only relevant to mate choice, but to other kinds of signalling, i.e. aposematism.

It might be said that vivid colours are, on average, more discriminable from typical backgrounds. This is probably true. Although detection and discrimination are certainly important in visual communication, they do not account for the super-threshold judgements made about objects already detected. This is necessary if one wishes to discuss signals. For this reason, we have avoided these concepts in our arguments. Our account rests on very different principles and provides a wholly different perspective; one which addresses preference and salience above the level of detection, rather than the thresholds that limit it.

Models of visual salience to humans typically include components that are akin to vividness (Niebur and Koch, 1996), and this kind of low-level prediction can be compared and contrasted by higher level theories based on asking subjects about the aesthetic value of a colour. Palmer and Schloss's (Palmer and Schloss, 2010) valence theory proposes that humans prefer colours associated with desirable objects, to those associated with decay, excrement and so forth. Given that decomposition and biological waste tend to be chemical mixtures, with low vividness, whereas pure materials tend to be more vivid it would be interesting to test whether vividness predicts colour preference as well as the valence (i.e. associated affect). At a more practical level, we can offer some assurance to field biologists that it is reasonable to generalise from our own colour perception to that of other animals, despite their physiological differences (Bennett et al., 1994).

Acknowledgements

We would like to thank Mary Stoddard, Almut Kelber, Justin Marshall and our reviewers for their comments, and Justin Marshall also for the spectra used in Fig. 1.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: L.W., D.O.; Methodology: L.W.; Software: L.W.; Formal analysis: L.W.; Writing - original draft: L.W., D.O.; Writing - review & editing: L.W., D.O.; Visualization: L.W.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data availability

The Munsell chip data are available from the Spectral Color Research Group, University of Eastern Finland at http://www.uef.fi/web/spectral/-spectral-database (Parkkinen et al., 1989). The code used to generate figures here has been made available, along with documentation and usage examples, as an open source project currently hosted at https://github.com/lucas-wilkins/lemonsauce.

Supplementary information

Supplementary information available online at
http://jeb.biologists.org/lookup/doi/10.1242/jeb.204487.supplemental

References

Allen, G. (1879). *The Colour Sense*. Wentworth Press.

Aronsson, M. and Gamberale-Stille, G. (2008). Domestic chicks primarily attend to colour, not pattern, when learning an aposematic coloration. *Anim. Behav.* **75**, 417-423. doi:10.1016/j.anbehav.2007.05.006

Bennett, A. T. D., Cuthill, I. C. and Norris, K. J. (1994). Sexual selection and the mismeasure of color. *Am. Nat.* **144**, 848-860. doi:10.1086/285711

Briscoe, A. D., Bybee, S. M., Bernard, G. D., Yuan, F., Sison-Mangus, M. P., Reed, R. D., Warren, A. D., Llorente-Bousquets, J. and Chiao, C.-C. (2010). Positive selection of a duplicated uv-sensitive visual pigment coincides with wing pigment evolution in heliconius butterflies. *Proc. Natl Acad. Sci. USA* **107**, 3628-3633. doi:10.1073/pnas.0910085107

Bybee, S. M., Yuan, F., Ramstetter, M. D., Llorente-Bousquets, J., Reed, R. D., Osorio, D. and Briscoe, A. D. (2011). Uv photoreceptors and uv-yellow wing pigments in heliconius butterflies allow a color signal to serve both mimicry and intraspecific communication. *Am. Nat.* **179**, 38-51. doi:10.1086/663192

Carroll, J., Murphy, C. J., Neitz, M., Ver Hoeve, J. N. and Neitz, J. (2001). Photopigment basis for dichromatic color vision in the horse. *J. Vis.* **1**, 2-2. doi:10.1167/1.2.2

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection*. John Murray.

Endler, J. A. (1990). On the measurement and classification of colour in studies of animal colour patterns. *Biol. J. Linn. Soc.* **41**, 315-352. doi:10.1111/j.1095-8312.1990.tb00839.x

Endler, J. A. and Basolo, A. L. (1998). Sensory ecology, receiver biases and sexual selection. *Trends Ecol. Evol.* **13**, 415-420. doi:10.1016/S0169-5347(98)01471-2
Endler, J. A. and McLellan, T. (1988). The processes of evolution: toward a newer synthesis. *Annu. Rev. Ecol. Syst.* **19**, 395-421. doi:10.1146/annurev.es.19.110188.002143
Fisher, R. A. (1930). *The Genetical Theory of Natural Selection: a Complete Variorum Edition*. Oxford University Press.
Franks, D. W., Ruxton, G. D. and Sherratt, T. N. (2009). Warning signals evolve to disengage batesian mimics. *Evolution* **63**, 256-267. doi:10.1111/j.1558-5646.2008.00509.x
Fuller, R., Houle, D. and Travis, J. (2005). Sensory bias as an explanation for the evolution of mate preferences. *Am. Nat.* **166**, 437-446. doi:10.1086/444443
Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Psychology Press.
Gittleman, J. L. and Harvey, P. H. (1980). Why are distasteful prey not cryptic? *Nature* **286**, 149-150. doi:10.1038/286149a0
Govardovskii, V. I., Fyhrquist, N., Reuter, T., Kuzmin, D. G. and Donner, K. (2000). In search of the visual pigment template. *Vis. Neurosci.* **17**, 509-528. doi:10.1017/S0952523800174036
Grafen, A. (1990). Biological signals as handicaps. *J. Theor. Biol.* **144**, 517-546. doi:10.1016/S0022-5193(05)80088-8
Guilford, T. and Dawkins, M. S. (1991). Receiver psychology and the evolution of animal signals. *Anim. Behav.* **42**, 1-14. doi:10.1016/S0003-3472(05)80600-1
Hamilton, W. and Zuk, M. (1982). Heritable true fitness and bright birds: a role for parasites? *Science* **218**, 384-387. doi:10.1126/science.7123238
Hart, N. S., Lisney, T. J., Marshall, N. J. and Collin, S. P. (2004). Multiple cone visual pigments and the potential for trichromatic colour vision in two species of elasmobranch. *J. Exp. Biol.* **207**, 4587-4594. doi:10.1242/jeb.01314
Hill, G. E. (1991). Plumage coloration is a sexually selected indicator of male quality. *Nature* **350**, 337. doi:10.1038/350337a0
Johnstone, R. A. (1995). Sexual selection, honest advertisement and the handicap principle: Reviewing the evidence. *Biol. Rev.* **70**, 1-65. doi:10.1111/j.1469-185X.1995.tb01439.x
Kelber, A., Vorobyev, M. and Osorio, D. (2003). Animal colour vision – behavioural tests and physiological concepts. *Biol. Rev.* **78**, 81-118. doi:10.1017/S1464793102005985
Kemp, D. J., Herberstein, M. E., Fleishman, L. J., Endler, J. A., Bennett, A. T., Dyer, A. G., Hart, N. S., Marshall, J. and Whiting, M. J. (2015). An integrative framework for the appraisal of coloration in nature. *Am. Nat.* **185**, 705-724. doi:10.1086/681021
Koenderink, J. (2010). *Color for the Sciences*. MIT Press.
Livingstone, M. and Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science* **240**, 740-749. doi:10.1126/science.3283936
Logvinenko, A. D. (2009). An object-color space. *J. Vis.* **9**, 5. doi:10.1167/9.11.5
Maloney, L. T. (1986). Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *J Opt Soc Am A* **3**, 1673-1683. doi:10.1364/JOSAA.3.001673
Marshall, J., Carleton, K. L. and Cronin, T. (2015). Colour vision in marine organisms. *Curr. Opin. Neurobiol.* **34**, 86-94. doi:10.1016/j.conb.2015.02.002
Maxwell, J. C. (1860). On the theory of compound colours, and the relations of the colours of the spectrum. *Phil. Trans. R. Soc. Lond.* **150**, 57-84. doi:10.1098/rstl.1860.0005
Munsell Color Co. (1950). *Munsell Book of Color*. Munsell Color Co.
Niebur, E. and Koch, C. (1996). Control of selective visual attention: modeling the 'where' pathway. *Neural Inf. Process. Syst.* **8**, 802-808.
Olson, V. A. and Owens, I. P. (1998). Costly sexual signals: are carotenoids rare, risky or required? *Trends Ecol. Evol.* **13**, 510-514. doi:10.1016/S0169-5347(98)01484-0
Olsson, P., Lind, O. and Kelber, A. (2017). Chromatic and achromatic vision: parameter choice and limitations for reliable model predictions. *Behav. Ecol.* **29**, 273-282. doi:10.1093/beheco/arx133
Osorio, D. and Vorobyev, M. (2008). A review of the evolution of animal colour vision and visual communication signals. *Vision Res.* **48**, 2042-2051. doi:10.1016/j.visres.2008.06.018
Osorio, D. and Bossomaier, T. R. (1992). Human cone-pigment spectral sensitivities and the reflectances of natural surfaces. *Biol. Cybern.* **67**, 217-222. doi:10.1007/BF00204394
Osorio, D. and Ham, A. D. (2002). Spectral reflectance and directional properties of structural coloration in bird plumage. *J. Exp. Biol.* **205**, 2017-2027.
Osorio, D. and Vorobyev, M. (2005). Photoreceptor spectral sensitivities in terrestrial animals: adaptations for luminance and colour vision. *Proc. R. Soc. B* **272**, 1745-1752. doi:10.1098/rspb.2005.3156
Palmer, S. E. and Schloss, K. B. (2010). An ecological valence theory of human color preference. *Proc. Natl Acad. Sci. USA* **107**, 8877-8882. doi:10.1073/pnas.0906172107
Parkkinen, J. P. S., Hallikainen, J. and Jaaskelainen, T. (1989). Characteristic spectra of Munsell colors. *JOSA A* **6**, 318-322. doi:10.1364/JOSAA.6.000318
Prum, R. O. (2012). Aesthetic evolution by mate choice: Darwin's really dangerous idea. *Phil. Trans. R. Soc. B* **367**, 2253-2265. doi:10.1098/rstb.2011.0285
Renoult, J. P., Kelber, A. and Schaefer, H. M. (2017). Colour spaces in ecology and evolutionary biology. *Biol. Rev.* **92**, 292-315. doi:10.1111/brv.12230
Rowe, C. and Guilford, T. (1996). Hidden colour aversions in domestic chicks triggered by pyrazine odours of insect warning displays. *Nature* **383**, 520. doi:10.1038/383520a0
Schrödinger, E. (1920). Theorie der Pigmente von grösster Leuchtkraft. *Ann. Phys.* **4**, 603-622. doi:10.1002/andp.19203671504
Schrödinger, E. (1944). *What Is Life?* Reprinted 2012. Cambridge University Press.
Siddiqi, A., Cronin, T. W., Loew, E. R., Vorobyev, M. and Summers, K. (2004). Interspecific and intraspecific views of color signals in the strawberry poison frog *Dendrobates pumilio*. *J. Exp. Biol.* **207**, 2471-2485. doi:10.1242/jeb.01047
Stockman, A. and Sharpe, L. T. (2000). The spectral sensitivities of the middle-and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Res.* **40**, 1711-1737. doi:10.1016/S0042-6989(00)00021-3
Tokunaga, R. and Logvinenko, A. D. (2010). Material and lighting dimensions of object colour. *Vision Res.* **50**, 1740-1747. doi:10.1016/j.visres.2010.05.030
Vorobyev, M., Gumbert, A., Kunze, J., Giurfa, M. and Menzel, R. (1997). Flowers through insect eyes. *Isr. J. Plant Sci.* **45**, 93-101. doi:10.1080/07929978.1997.10676676
Vorobyev, M. (2003). Coloured oil droplets enhance colour discrimination. *Proc. R. Soc. B* **270**, 1255-1261. doi:10.1098/rspb.2003.2381
Wallace, A. R. (1878). Tropical nature, and other essays. Macmillan & Co.
Wyszecki, G. and Stiles, W. S. (2000). *Color Science: Concepts and Methods, Quantitative Data and Formulae*. Wiley-Interscience.
Zahavi, A. (1975). Mate selection – a selection for a handicap. *J. Theor. Biol.* **53**, 205-214. doi:10.1016/0022-5193(75)90111-3